orrick

November 18, 2024

Sent via ECF

Hon. Ona T. Wang
Daniel Patrick Moynihan
United States Courthouse
500 Pearl Street
New York, NY 10007-1312

Orrick, Herrington & Sutcliffe LLP

The Orrick Building
405 Howard Street
San Francisco, CA 94105-2669

+1 415 773 5700
orrick.com

Annette L. Hurst

E ahurst@orrick.com
D +1 415 773 4585
F +1 415 773 5759

RE: *Authors Guild et al. v. OpenAI Inc. et al.*, No. 1:23-cv-8292 and
       Alter et al. v. OpenAI Inc. et al., No. 1:23-cv-10211

Dear Judge Wang:

Microsoft respectfully requests a pre-motion conference regarding the Class Plaintiffs' refusal to produce documents relevant to the first and fourth fair use factors and the defense of substantial non-infringing uses. Specifically, as to the first factor, Plaintiffs are refusing to produce the documentation reflecting their *own* usage of ChatGPT. As to the fourth factor, Plaintiffs concede the relevance of the requested evidence but refuse to collect responsive documents within their possession or their practical ability to obtain.

Evidence regarding Plaintiffs' Usage of Generative AI. As part of its fair use defense, Microsoft must provide evidence to support the first factor—the purpose and character of the allegedly infringing use. 17 U.S.C. § 107(1). On the first factor, the primary inquiry concerns "whether the new work, '. . . merely supersede[s] the objects of the original creation, . . . or instead adds something new, with a further purpose . . . . [I]t asks, in other words, whether and to what extent the new work is 'transformative.'" *Authors Guild v. Google, Inc.*, 804 F.3d 202, 214-15 (2d Cir. 2015) (quoting *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 578-79 (1994)). How people use the generative AI products at issue is directly relevant to the analysis of the purpose of the use. In addition, Microsoft is defending Plaintiffs' claims for secondary liability of contributory copyright infringement on the ground that there are substantial noninfringing uses. *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 417 (1984) (holding that certain videocassette recorders were "capable of substantial noninfringing uses" due to evidence of "[p]rivate, noncommercial time-shifting in the home"); *Cartoon Network LP, LLLP v. CSC Holdings, Inc.*, 536 F.3d 121, 133 (2d Cir. 2008) ("asking whether copying copyrighted material is only 'incidental' to a given technology is akin to asking whether that technology has 'commercially significant noninfringing uses,' another inquiry the *Sony* Court found relevant to whether imposing contributory liability was just.") (quoting *Sony*, 464 U.S. at 442). Beneficial noninfringing uses are also relevant to balancing the alleged harm to the works in connection

with the fourth factor of fair use. *See Google LLC v. Oracle America, Inc.*, 593 U.S. 1, 35 (2021) ("we must take into account the public benefits the copying will likely produce").

Accordingly, Microsoft propounded discovery requests seeking information about how the Plaintiffs themselves are using ChatGPT. Ex. A (Authors Guild's Resps. & Objs. to Microsoft's First Set of Requests for Production) at 20 (Request No. 23); Ex. B. (Individual Plaintiffs' Resps. & Objs. to Microsoft's First Set of Requests for Production) at 21 (same). Plaintiffs have refused to provide documents responsive to Microsoft's request for *all* non-privileged prompt attempts. This motion is directed only to general usage of the tool and is not directed to the prompts used to produce evidence quoted in the Complaint that was addressed in a prior motion.

Evidence of Plaintiffs' use of ChatGPT is directly relevant to Microsoft's fair use and substantial non-infringing uses defenses. Plaintiffs' interactions with ChatGPT (including their prompts) will show the transformative and beneficial nature of large language models far beyond the cherry-picked uses alleged in Plaintiffs' Consolidated Complaint. Plaintiffs' use of large language models to assist with research and writing demonstrates specific pro-copyright benefits to these transformative and noninfringing uses. This is exactly the type of evidence this Court has previously compelled in other copyright cases. *See Viacom Int'l Inc. v. Youtube Inc.*, 253 F.R.D. 256, 261-62 (S.D.N.Y. 2008) (granting motion to compel evidence of substantial non-infringing uses).

Plaintiffs have not argued (let alone demonstrated) any burden whatsoever with producing these documents. They just want to improperly restrict the scope of the evidence to that which they choose to produce rather than what Defendant Microsoft is requesting. Microsoft respectfully requests the Court compel Plaintiffs to produce all documents relating to Plaintiffs' use and prompting of ChatGPT in response to Author's Guild Request No. 23 and Individual Plaintiffs Request No. 23.

Fourth Factor Fair Use Evidence. The documents requested related to these topics are essential to assessing the fourth factor of fair use: "the effect of the use upon the potential market for or value of the copyrighted work." 17 U.S.C. § 107(4). The law is clear: as a Defendant, Microsoft has the burden of presenting evidence on the fourth factor—one of the most important elements of fair use. *See Harper & Row, Publ'rs. v. Nation Enters.*, 471 U.S. 539, 566 (1985); *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 590 (1994) (burden on defendant to come forward with evidence); *Authors Guild v. Google, Inc.*, 804 F.3d 202, 223 (2d Cir. 2015) ("the fourth factor is of great importance in making a fair use assessment"). To that end, Microsoft has served specific, tailored requests on Plaintiffs such to assess the market for (including the potential market) and value of the copyrighted works-in-suit. *See* Ex. A & Ex. B: Request No. 12 (valuation of works before and after release of OpenAI's LLMs), Request No. 10 (documents concerning the market for the works-in-suit—before and after release of OpenAI's LLMs)—including gross revenues, the consumer base, and other market performance metrics as well as studies), and Request No. 11 (same, for derivative markets).

Plaintiffs have conceded the relevance of what was requested and have therefore agreed to produce documents and communications reflecting any valuation or market analysis of the Class Works, including, but not limited to, revenue reports, market performance reports in both traditional and potential markets, and market and valuation research and studies. Despite this agreement, Plaintiffs have improperly restricted the scope of their collection for such responsive material. The remaining problem is twofold: (1) first, Author's Guild is refusing to collect and produce all information in its possession; and (2) the Individual Plaintiffs have not agreed to ask those from whom they have a practical ability to obtain the documents—literary agents and publishers—for this information as well.

Plaintiffs concede that the relevant evidence here in considering harm to the "market" for the copyrighted work must encompass *all* traditional, reasonable, or likely to be developed markets in which the Plaintiffs are or could be selling their works. *Am. Geophysical Union v. Texaco Inc.*, 60 F.3d 913, 930 (2d Cir. 1994); *see also Google LLC v. Oracle America, Inc.*, 593 U.S. 1, 35-36 (2021) ("[A] potential loss of revenue is not the whole story. We here must consider not just the amount but also the source of the loss."). To weigh against fair use, "[t]here must be a meaningful or significant effect 'upon the potential market for or value of the copyrighted work.'" *Authors Guild*, 804 F.3d at 224 (quoting 17 U.S.C. § 107). In order to determine the existence of any such effect, Microsoft must first be able to understand the *full* scope of the market, potential market, and value of the copyrighted works. Thus, Microsoft propounded Requests which seek evidence directly related to the market for the copyrighted works-in-suit and their derivatives. *See* Ex. A at 12-13; Ex. B at 10-13.

First, The Authors Guild's takes the position that it must only collect the information in its possession related to the works of one author, Mignon Eberhart, because those are the only works it is asserting. This is an indefensible position. The Author's Guild is the preeminent representative of authors in the publishing industry. It regularly collects and publishes data regarding the performance of that industry, finances and compensation for authors, and marketing of books. The fact that Author's Guild is only asserting copyright directly on behalf of one author does not excuse it from its obligation to search for highly relevant market data within its possession. The Authors Guild should be ordered to produce all fourth factor evidence in its possession without regard to the works of Mignon Eberhart.

Second, with respect to the Individual Plaintiffs, given the obvious, and agreed-upon, relevance of this information, Plaintiffs should collect not only the data in their possession, but should also request data regarding their works from those with whom they work in publishing the works at issue. All of the Plaintiffs have commercial publishers and most (if not all) of them also have literary agents. Requesting relevant market and valuation and data from their publishers and agents is far more efficient that forcing Microsoft to seek it all through third-party subpoenas. Plaintiffs and Microsoft continue to meet and confer on this point, but if the Individual Plaintiffs refuse to make such requests from their agents, they should be ordered to do so. *See Bank of New York v. Meridien Biao Bank Tanz.*, 171 F.R.D. 135, 146 (S.D.N.Y. 1997) (granting motion to compel and finding that the practical ability test extends to "obtain[ing] documents from a non-party"); *see also Mirlis v. Greer*, 80 F.4th 377, 382 (2d Cir. 2023).

Respectfully submitted,

/s/ Annette L. Hurst

Annette L. Hurst